Chromatin: **The nucleosome unwrapped**

J. Widom

The structure of the nucleosome core particle has been determined by X-ray crystallography at 2.8 Å resolution. The structure has several significant surprises, and provides important new insights into the structure and function of chromatin.

Address: Department of Biochemistry, Molecular Biology and Cell Biology, and Department of Chemistry, Northwestern University, Evanston, Illinois 60208-3500, USA.
E-mail: j-widom@nwu.edu

Current Biology 1997, 7:R653–R655
http://biomednet.com/elecref/09609822007R0653

The nucleosome core particle is the fundamental subunit of chromatin [1,2]. It consists of two molecules each of the four 'core histone' proteins, H2A, H2B, H3 and H4, and a 147 base pair stretch of DNA. The lowest level of chromatin organization consists of a repeated array of nucleosome core particles, separated by variable lengths of 'linker DNA'. In many, but not all, cases, each core particle plus linker DNA is associated with one molecule of a fifth 'linker' histone, H1. The complex of the core particle plus its linker DNA and H1 (if present) is called a nucleosome.

Nucleosomes are of intense interest, in part because of their importance in chromatin function. In eukaryotes, the substrates for the essential biological processes of transcription, replication, recombination, DNA repair and cell division are chromatin, not naked DNA. Nucleosomes play both active and passive roles, in both positive and negative gene regulation. The structural properties of nucleosomes are also remarkable. DNA in nucleosomes is wrapped around the histone octamer in ~1¾ superhelical turns, and so is bent exceptionally tightly in comparison to the length scale of its flexibility.

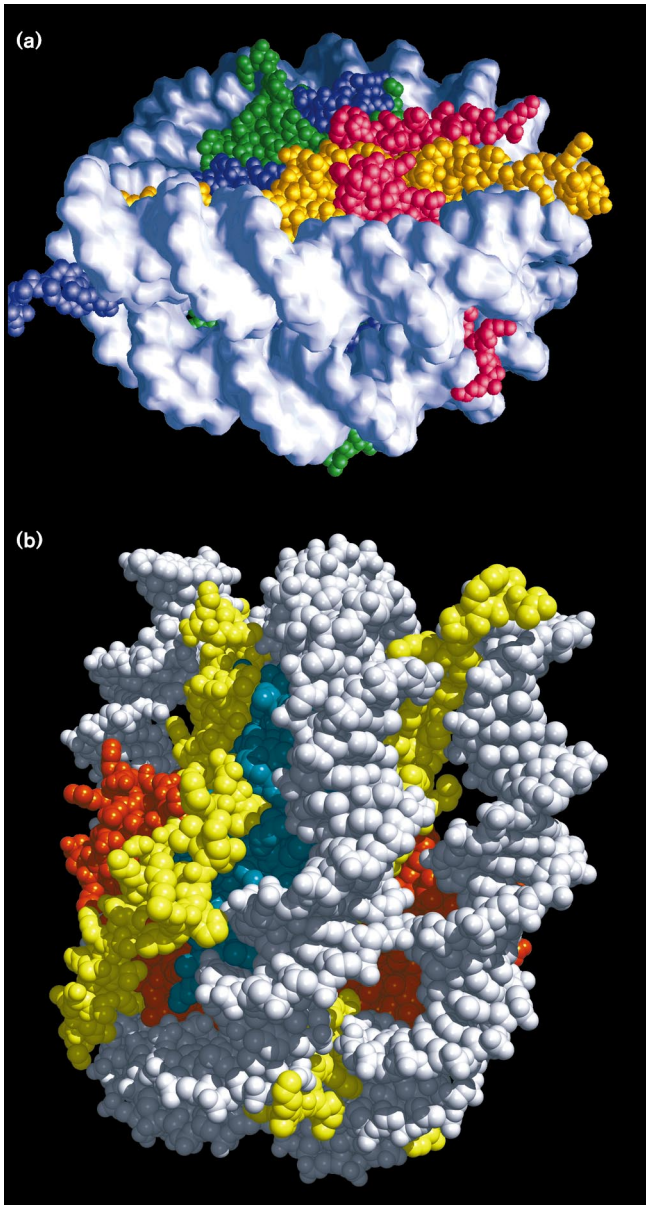
The structure of the nucleosome core particle was solved at 7 Å resolution in 1984 [3], and that of the histone octamer on its own was solved at 3.1 Å resolution in 1991 [4]. These structures revealed a basic tripartite assembly of the octamer, reflecting its two H2A/H2B heterodimer and one $H3_2H4_2$ tetramer components. The tetramer itself is a stable complex of two H3/H4 heterodimers, which have a 'handshake' interlocking protein fold very similar in structure to that of the H2A/H2B heterodimers. Each histone dimer has a pseudo-two-fold symmetry. This protein architectural motif is now referred to as the 'histone fold', which has subsequently been seen in two 'TAF' subunits of the *Drosophila* transcription factor TFIID, and in archaebacterial histone-like proteins. The octamer has an overall pseudo-two-fold (dyad) symmetry. A prominent feature of the octamer surface is a positively charged superhelical ramp, important for the DNA organization [3].

Now, in a major accomplishment that is the culmination of 13 years of effort, the structure of the nucleosome core particle has been revealed by X-ray crystallography in atomic detail (Figure 1) [5]. This required the use of reconstituted nucleosomes prepared from homogeneous recombinant histones, and a particular DNA sequence that had been found by happenstance to form high quality nucleosome crystals, together with a high-intensity synchrotron X-ray source. In addition to providing higher resolution views of the protein core itself than were previously available, this new structure provides detailed views of the DNA and the protein–DNA interface; it offers tantalizing glimpses of the histone tail domains; and it adds support to solenoidal models for the 30 nm chromatin fiber, the next higher level of chromatin structure [2].

The structure contains many real surprises, starting with the length of DNA in the particle. The particles were crystallized with a symmetric 146 base pair DNA fragment; rather than sitting symmetrically about the octamer's pseudo-dyad axis, however, the DNA is positioned asymmetrically, with a 73 base pair half, a 72 base pair half, and one base pair located on the pseudo-dyad axis. Previous nuclease digestion studies had established that the minor groove of the DNA faced outward at the particle dyad, but it was generally assumed that the pseudo-two-fold axis passed between base pairs. A second surprise is that, probably in order to accommodate the requirements of an inter-particle contact in the crystal, a 12 base pair region in the 72 base pair half is substantially overwound and stretched, so as to return after 12 base pairs to a location similar to that of the corresponding 13 base pairs in the 73 base pair half.

Overall, the DNA is wrapped around the histone octamer in ~1.65 superhelical turns, with a pitch of ~26.5 Å and a radius of ~42.5 Å (at the DNA helix axis), although irregularities in the wrapping cause these to vary significantly throughout the particle. The helical twist averaged over the full length is ~10.2 base pairs per turn, as opposed to the ~10.5–10.6 base pairs per turn characteristic of DNA in solution. The twist of 10.2 base pairs per turn matches periodicities found in sequence motifs in isolated nucleosomes and in whole genomes, implying that genomes have evolved in part to aid in their own nucleosomal packaging [6]. In more detail, however, the helical twist

varies from place to place along the DNA — even ignoring the distortions caused by the shorter length of the 72 base pair arm — in a way that had not been anticipated. Bending is distributed along the length of the DNA and includes bends into both the major and minor grooves. Eight sites — at superhelix locations ± 1, 2, 4 and 5 in their numbering system — that involve bends into the major grooves appear particularly deformed. The histone fold domains organize the central 121 base pairs of DNA, with the additional 13 base pairs at each end organized by an amino-terminal α helical extension to the histone fold of H3 and preceding residues from the tail domain. Each histone dimer contributes three main DNA-binding motifs, in two types, referred to as L1L2 and α1α1. These occur in the order L1L2, α1α1, L1L2, spaced ~10 base pairs apart.



DNA binding is primarily to the sugar phosphate backbone over the short stretches of each DNA helical turn, where the minor groove — and hence the DNA backbone — faces in towards the octamer surface. Contacts between the histones and DNA are primarily nonspecific. They include extensive salt bridges and hydrogen bonds to the phosphate groups, contributed by both main-chain and side-chain groups; extensive nonpolar contacts with the DNA sugar groups; and electrostatic interactions of the positively charged amino termini of α helices with DNA phosphates. However, there are also a number of base-specific contacts, including nonpolar contact of the 5-methyl group of thymidine in the major groove.

Two additional modes of histone–DNA interaction are particularly striking. First, an arginine side chain is inserted into the minor groove every time it faces inward to the histone surface. In most cases, the arginine is held by additional bonds to protein functional groups so as to prevent it from penetrating deeply into the groove and making base-specific contacts. Second, several of the histone tails act together to bracket turns of DNA, passing over and between the DNA gyres. Each of the core histones has an ~10–40 residue, highly positively charged amino-terminal region; and histones H2A and H3 have shorter but analogous domains at their carboxyl termini as well. These domains are referred to as 'tails' because they are known to be highly extended and mobile.

These histone tail domains are of great interest for several reasons. They are the sites of numerous post-translational modifications that are known to be essential in chromatin function, including a diverse set of specific acetylations that have been the subject of intense recent interest in the field of gene regulation [7]. The tails also mediate interactions with a variety of other proteins involved in gene regulation [8]. The amino-terminal domains of H2B and H3 pass between the gyres of DNA, threading through closely apposed minor grooves and extending out from the particle.

Figure 1

The nucleosome core particle at high resolution. **(a)** A combined space-filling and surface representation. DNA is white; histones H2A, H2B, H3 and H4 are colored yellow, red, blue and green, respectively. The altered DNA helical twist brings the major and minor grooves on neighboring gyres of DNA into striking alignment. The amino-terminal tails of H3 and H2B protrude through closely apposed minor grooves out from between the DNA gyres; the amino-terminal tails of H2A pass over the DNA along a minor groove. The upper face of the nucleosome is dominated by an H2A/H2B heterodimer; these residues could be important in an interaction with the amino-terminal domain of H4 from another nucleosome, neighboring in space in the chromatin solenoid. **(b)** A space-filling representation. DNA is white; the histone fold parts of the $H3_2H4_2$ tetramer and the H2A/H2B dimers are turquoise and orange, respectively. The tails and extensions are yellow. The tails and extensions of the basic histone fold dominate much of the nucleosome surface. (Figures kindly provided by K. Luger and T. Richmond.)

The amino-terminal domains of H2A pass over the DNA along a minor groove.

Among other roles, these tails seem likely to influence the energetics and dynamics of DNA site exposure. Site exposure of nucleosomal DNA is a mechanism whereby gene regulatory proteins and polymerases gain access to their nucleosomal DNA substrates, making possible the cooperative invasion of nucleosomes [9,10]. Measured equilibrium constants for site exposure decrease more or less progressively as one moves inward from an end into the middle of the nucleosome. Such behavior, and the structure of the nucleosome itself, are consistent with a picture in which DNA spontaneously and transiently uncoils inward starting from an end. The structure shows the DNA wrapped on the histone surface as making contacts ('bonds') in a small patch, approximately every 10 base pairs, each time the phosphodiester backbone (minor groove) faces inward toward the octamer. Thus, uncoiling would naturally proceed stepwise, with an incremental increase in energetic cost — decreased equilibrium constant for site exposure — associated with each additional 10 base pair segment uncoiled. Post-translational modifications of the tail domains would be expected to influence the equilibrium constants for site exposure — in other words, to influence the time-averaged accessibility of the nucleosomal DNA.

The structure also leads one to re-examine the issues of DNA sequence-directed nucleosome positioning. Positioning achieves equilibrium, so the sites of preferential positioning will be those having minimum free energy. The net free energy for any particular position will reflect favorable contributions from the set of all the bonds that are formed — including all intermolecular, intramolecular and solvent bonds — minus the free energy cost of deforming the protein, the DNA and the solvent away from their starting (uncomplexed) conformations into their core particle conformations. While the structures of the L1L*2*–DNA and $\alpha 1\alpha 1$–DNA interactions are relatively well conserved throughout the core particle, they differ in detail, probably from the necessity of accommodating differing local DNA sequences. We already know that the DNA changes structure upon nucleosome formation, and it is most likely that the histone octamer and the solvent do too. Thus the detailed equilibrium structures of the nucleosome will vary in detail with the DNA sequence.

The structure of the core particle shows myriad opportunities for particular DNA sequences to influence the number and strength of bonds that are made in the complex, as well as the energetic cost of changing the bondedness of the separated partners. The structure also shows myriad locations where specific DNA bends or twists may be optimal. Thus, the DNA sequence can also contribute significantly to the energetics through its effect on the mechanical work involved in changing the position-dependent DNA bending and twist, which reflect contributions from four sources: static bending, the bendability (bending force constant), static twist, and the twistability (twisting force constant). The results of a parallel study on the structure of a core particle containing a different DNA sequence, which is in progress in the Richmond laboratory [5], will be particularly interesting.

A final great surprise is a completely unanticipated strong interparticle protein–protein contact between the amino-terminal tail of histone H4 on one nucleosome and an evolutionarily conserved patch on an H2A/H2B heterodimer on an adjacent nucleosome. A similar contact — albeit with reversed geometry, made possible by the flexibility of the tail and the nature of the binding interface — seems likely to serve as an important stabilizing element in the packing of adjacent nucleosomes in the 30 nm chromatin fiber [2]. This very important contribution will serve as a focus for much of the work in this field for years to come.

References

1. van Holde KE: *Chromatin.* New York: Springer-Verlag; 1989.
2. Widom J: **Toward a unified model of chromatin folding.** *Annu Rev Biophys Biophys Chem* 1989, **18:**365-395.
3. Richmond TJ, Finch JT, Rushton B, Rhodes D, Klug A: **Structure of the nucleosome core particle at 7 Å resolution.** *Nature* 1984, **311:**532-537.
4. Arents G, Burlingame RW, Wang B-C, Love WE, Moudrianakis EN: **The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix.** *Proc Natl Acad Sci USA* 1991, **88:**10148-10152.
5. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ: **Structure of the nucleosome core particle at 2.8 Å resolution.** *Nature* 1997, **389:**251-260.
6. Lowary PT, Widom J: **Nucleosome packaging and nucleosome positioning of genomic DNA.** *Proc Natl Acad Sci USA* 1997, **94:**1183-1188.
7. Wade PA, Wolffe AP: **Chromatin: histone acetyltransferases in control.** *Curr Biol* 1997, **7:**R82-R84.
8. Hecht A, Laroche T, Strahl-Bolsinger S, Gasser SM, Grunstein M: **Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast.** *Cell* 1995, **80:**583-592.
9. Polach KJ, Widom J: **A model for the cooperative binding of eukaryotic regulatory proteins to nucleosomal target sites.** *J Mol Biol* 1996, **258:**800-812.
10. Protacio RU, Widom J: **Nucleosome transcription studied in a real-time synchronous system: test of the lexosome model and direct determination of effects due to histone octamer.** *J Mol Biol* 1996, **256:**458-472.